PROSPECTIVE TEACHERS' CONCEPTIONS OF PROOF

Zulfiye Zeybek
Indiana University
zzeybek@indiana.edu

What types of mathematical justifications do pre-service elementary teachers find convincing? To investigate this question, a task-based interview which was designed to elicit arguments of what students find convincing was administered to two female students who were enrolled in a geometry course at a large Midwestern university. These arguments were categorized according to the proof schemes crafted by analyzing different studies dealing with proof. A qualitative analysis of the data revealed that the two pre-service elementary teachers (PSTs) who were interviewed have difficulties in following or constructing formally presented deductive arguments and in understanding how deductive arguments differ from inductive arguments. They also held explicit misconceptions about proving (or disproving) statements such as: "a couple of examples constitute proof" or "one counterexample is not sufficient to disprove a statement."

Keywords: Inductive Argument; Deductive Argument; Proofs; Preservice Elementary Teachers

Introduction

It is difficult to overstate the importance of proofs in mathematics. If you have a conjecture, the only way that you can be completely sure that it is true is by presenting a valid mathematical proof. However, the mathematics education community worldwide is facing the challenge of improving students' abilities to prove and to reason mathematically at all grade levels. Despite the fundamental role that proof and refutation play in mathematical inquiry (Lakatos, 1976) and the growing appreciation of the importance of these concepts in students' mathematical education (Hanna, 2000; Reid, 2002), students hold various misconceptions not only about proof, but also about refutation (Chazan, 1993; Simon & Blume, 1996).

Several studies have reported that formal deduction among students who have studied secondary school geometry is nearly absent (Burger & Shaughnessy, 1986; Dreyfus, 1999; Chazan, 1993). Many students accept inductive arguments as valid mathematical proof (Martin & Harel, 1989; Chazan, 1993) or they fail to recognize that using a larger set of examples still does not constitute proof (Knuth, Choppin, & Bieda, 2009). In addition, students have difficulty in understanding that a valid proof confers the universal truth of a general statement; thus mathematical proof requires no further empirical verification (Fischbein, 1982; Martin & Harel, 1989; Chazan, 1993). Some students believe that counterexamples do not really refute; instead they tend to treat valid counterexamples to general statements as exceptions that do not really affect the truth of the statements (Balacheff, 1988). Similarly, Simon and Blume (1996) show that many students think that giving one example is not enough to refute an argument.

Despite students' current lack of knowledge, as well as interest, in proof and proving, the topic is central to mathematics, so it should be a key component of mathematics education (Bell, 1976; Hanna, 2000; Martin & Harel, 1989). Not only is proof at the heart of mathematical practice, it is an essential tool for promoting mathematical understanding (Martin & Harel, 1989; Hanna, 2000; Knuth, 2002). Stylianides (2007) has shown that young children can make legitimate mathematical arguments and even formal arguments that count as proof. He claims that proof should be part of students' mathematical experiences even in early elementary grades (Stylianides, 2007). Similarly, Harel and Sowder (1998) argue that instructional activities that educate students to reason mathematically about situations are crucial to students' mathematical development, and that these activities must begin at an early age. Thus, calls for improvement in mathematics education in the U.S. have increasingly emphasized the importance of proof and reasoning by recommending that reasoning and proof should be a part of the mathematics curriculum at all levels from pre-kindergarten through grade 12 (NCTM, 2000).

The purpose of this paper is to describe pre-service elementary teachers' attempts to construct proofs and also to examine different arguments regarding proofs in order to better understand their conceptions of proof. The following research questions guided the study:

- How do pre-service elementary teachers support claims, warrants, and backings as elements of their argumentations?
- What are pre-service elementary teachers' conceptions of proof?

The Framework

Various studies that include a description of proof schema ideas at both pre-college and college levels are evaluated with an effort to craft the framework used in this study (see Table 1). Hanna (1989) argues that proofs can have different degrees of validity and still gain the same degree of acceptance. To document different types of mathematical justifications attempted by each participant, mathematical arguments constructed to examine pre-service teachers' conception of proof are assessed according to a hierarchy of levels of mathematical justifications explained in the framework.

While many studies have focused primarily on distinctions between the inductive and deductive justifications (Chazan, 1993; Martin & Harel, 1989; Morris, 2002), some researchers have posed questions such as: What might make one example or empirical justification stronger than another? Or can all mathematical arguments be categorized as inductive or deductive? As a result, they have divided inductive and deductive justifications into further subcategories (Balacheff, 1988; Harel & Sowder, 2007; Simon & Blume, 1996; Quinn, 2009) and proposed another type of justification along with inductive and deductive justifications (Harel & Sowder, 1998; Simon, 1996). The framework crafted for this study focuses not only on the distinction between the inductive and deductive justifications, but also further subdivides those categories as well as includes the justifications that are neither inherently inductive nor deductive.

Many researchers define several stages or levels in which students' reasoning skills vary in terms of the justifications they are able to produce (Bell, 1976; Simon & Blume, 1996; Quinn, 2009). Harel and Sowder (1998, 2007) categorize those levels into three classes in their taxonomy—the external conviction proof scheme class, the empirical proof scheme class, and the deductive proof scheme class—with some subschema for each class. Similarly, Balacheff (1988) describes two main categories—pragmatic justifications and conceptual justifications—which play complementary roles. The first justification type (pragmatic justification) is divided into three subcategories: naïve empiricism (justification by a few random examples), crucial experiment (justification by carefully selected examples), and generic examples (justification by an example representing salient characteristics of a whole class of cases) and the second justification type (conceptual justification) into two subcategories as "thought experiment" and "symbolic calculation." Weber and Alcock (2004), on the other hand, focus only on deductive reasoning and divide deductive justifications as syntactic proof scheme (manipulating correctly stated definitions and facts in a logically valid way) and semantic proof scheme (use instantiations of the mathematical objects to which the statement applies to suggest and guide the formal inferences) which aligns with Hanna's (2000) distinction of proofs that prove and proofs that also explain.

The framework used in this study summarizes the proof schemes explained above, by merging those categories—from external to analytic—along with different levels in which provers demonstrate different level of mathematical justifications. The framework outlines various strong background work, thus, provides a powerful as well as useful tool for an analytical assessment of PSTs' conceptions.

Table 1: Types of Mathematical Justifications

<table>
<tr><td colspan="2">Level 0</td><td>Responses that do not address justification.</td></tr>
<tr><td rowspan="2">EXTERNAL</td><td rowspan="2">Level 1</td><td>Appeals to external authority</td></tr>
<tr><td>Authoritarian proof : depends on an authority such as a teacher or a book
Ritual proof : depends on the appearance of the argument
Non-referential symbolic proof: depends on some symbolic manipulation, often without reference to the symbols’ meaning</td></tr>
<tr><td colspan="2">Level 2</td><td>Naïve reasoning, usually with incorrect conclusions. Although provers use some deduction, the arguments start with an analogy or with something that provers remember hearing, often incorrectly. Provers generally reach an incorrect conclusion or, if they reach a correct conclusion, they have used the wrong assumptions.</td></tr>
<tr><td rowspan="3">EMPIRICAL</td><td rowspan="2">Inductive Frame</td><td>Level 3A: Naïve Empiricism: an assertion is valid from a small number.</td></tr>
<tr><td>Level 3B: Crucial Empiricism deals more explicitly with the question of generalization by examining a case that is not very particular. If the assertion holds in that case, it is validated.</td></tr>
<tr><td>Rudimentary Transformational Frame</td><td>Level 3C: Perceptual Proof: Provers make inferences that are based on rudimentary mental images that are not fully supported by deduction.</td></tr>
<tr><td colspan="3">Deductive frame expressed in terms of particular instances</td></tr>
<tr><td>TRANSFORMATIONAL</td><td>Level 4</td><td>Generic Example: Deductive justification that is expressed in terms of a particular instance (examples might be used to generalize the rules, but unlike an empirical proof scheme, the general rules are predicted based on the inference rules.) Simon (1996) defines transformational proof schemes an enactment of an operation (or set of operations) on an object (or set of objects) that allows one to envision the transformations that these objects undergo and the set of results of these operations.</td></tr>
<tr><td colspan="3">Deductive frame that is independent of particular instances</td></tr>
<tr><td>ANALYTIC</td><td>Level 5</td><td>Syntactic: a verification of a statement is evaluated according to ritualistic features.
Semantic/Conceptual: a judgment is made according to causality and purpose of argument.</td></tr>
</table>

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.

Method

Participants

Two pre-service elementary teachers, Sara and Dacey (pseudonyms), volunteered to participate in this study. Both of the participants had enrolled in a Geometry-content course designed for elementary majors at a large Midwestern university. Both participants satisfied the course requirements and passed the course with a grade of B or above.

Data Sources

The study uses a qualitative approach, mainly participant classroom observation and task-based interviews to investigate pre-service elementary teachers' conceptions of proof. Every class in which the participants were enrolled was audiotaped and notes were taken by the author of this study.

Each participant was interviewed individually for about an hour in a semi-structured manner using an interview script consisting of three phases. The interviews took place near the end of the semester, in each case. Thus, the interviewees were expected to have learned most of the course topics and have had some practice justifying different statements by the time the interviews took place.

Phase 1. During this phase, each student was presented the written tasks A, B, and C, described below. The PSTs were asked to explain in their own words what the statements said, and to decide whether the statements were always, sometimes, or never true and how they would know. And then, they were asked to produce a justification in cases where they believed the statements to be true.

Phase 2. After letting the participants try to justify the statements by themselves first, the participants were presented with four brief arguments for both Task A and B, varying in terms of level of justification, one after the other, and asked to think out loud as they read each one, to judge the correctness, and to say to what extent each argument was convincing.

Phase 3. Having seen and thought about all four arguments, one after the other, the students were provided "Always," "Sometimes," "Never" cards and asked to assign the appropriate card to each argument presented. For instance, if the participants thought that the conclusion derived from one of the arguments would always hold true then they needed to put an "Always" card on the argument.

The data collected consisted of the audiotaped interviews, the interviewer's notes and the students' work on the "proof" sheets provided during the interview.

Interview Tasks

The interview tasks were designed to provide, first, an indication of pre-service elementary teachers' competence in constructing proofs, and then, an overview of their views as to what constituted a proof. The interview tasks included three types of items (from familiar to unfamiliar) to probe pre-service teachers' views of proof from a variety of standpoints.

Task A. This task was adopted from the course textbook. Thus, it was expected that by this time of the year, the interviewees were familiar with it and could reproduce the proof on their own. The task appears on the sheet presented to the participants as follows:

> A kite is a quadrilateral with two distinct pairs of adjacent sides that are equal in length. Given the definition, justify whether or not the following statement is true. "*In a kite, one pair of opposite angles is congruent*."

Task B. The same structure as in Task A was used to construct Task B. This task was adopted from Chazan (1993), but it was modified such that four arguments, varying in terms of level of justification, were added to present to the participants. The task appears as follows:

> Justify whether or not the statement is true: "*In any triangle, a segment joining the midpoints of any two sides will be parallel to the third side*."

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.

Task C. Task C was a non-familiar case to the participants. Thus, it was expected that this task might be challenging for them. This task was adopted from Simon and Blume (1996). The task appears as follows: Find the area of the shape below.

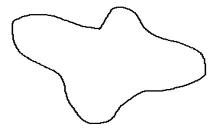

Figure 1

Arguments for interview tasks. The theoretical framework explained before (see table 1) governed the choice of arguments included in both of task A and B. Arguments for task A and B were characterized as empirical, subdivided as Naïve Empiricist with a small number of cases (Arguments 1 labeled as Level 3A) and Crucial Empiricism with an extended number of cases including non-particular cases (Arguments 2 labeled as Level 3B), argument requiring concrete demonstration or explanation written in everyday style (Arguments 3 labeled as Level 4), and a deductive proof, written in a formal style (two-column) (Arguments 4 labeled as Level 5). For task C, the following method was presented to the participants:

> "*If you take a piece of string and measure the whole outside of the area and then pull that into a shape like a rectangle, you can easily calculate the area of the figure*." Justify whether or not the above method will work to find the area of the figure.

Results

Sara's Proof Scheme

When Sara was presented task A, she attempted to dissect the kite into two triangles in order to use triangle congruency to justify the statement. However, because she drew a diagonal that produced two isosceles triangles instead of a diagonal that produced two congruent triangles, she failed to proceed from there to justify the statement. Even though, she started to use some deduction such as congruent triangles that she remembered hearing from her class, she failed to reproduce it correctly. Similarly, she attempted to use what she learned about parallel lines in her class to justify the statement in task B. However, she failed to proceed from using her previous knowledge to construct a justification. Thus, her proof scheme was coded as Level 2 (Naïve Reasoning) for both task A and B.

Even though Sara failed to construct a proof, she correctly distinguished the deductive arguments from the inductive arguments when she was presented the arguments for both task A and B. Sara understands that a couple of examples do not qualify as proof. She was aware that the conclusion that was arrived at from direct measurements of specific cases was approximate and that the generalization which was arrived at without examining every possible case might be highly probable but not certain. Additionally, Sara understands the role of justification in mathematics: that is, to provide an argument that holds for every case. She knows that providing examples will hold only for those specific examples and she chooses "Sometimes" for argument 1 and 2 and "Always" for argument 3 and 4. Thus, her proof scheme was coded as Level 5 in phase 2 and 3 for both tasks.

Dacey's Proof Scheme

Dacey, on the other hand, did not attempt to reproduce the proof she learned in her class for task A. Instead, she tried to justify the statement by saying that if two sides are equal, then the angles between them are going to be equal since where those sides will meet will be the same. Dacey did not attempt to provide examples nor attempted to use logical deduction to justify the statement. Rather, her attempt to prove the statement in Task A was driven by her perceptual observation of the figure provided to her. Thus, her proof scheme for this task was coded as Level 3C based on the framework. When Dacey was presented task B and asked to decide whether or not the statement was correct, she quickly concluded that the statement was correct, because, as she explained, the instructor recently showed the same statement

and justified why it was correct in the class. However, because Dacey did not understand why the statement was correct or how to justify that it was correct in class, she failed to reproduce the proof. Dacey remembered that the proof included corresponding angles, but she could not proceed from using corresponding angles to conclude that the statement is true. Her response for this task was coded as Level 2.

Dacey found arguments 3, 2 and 1 more convincing than argument 4. She claimed that seeing actual measurements or illustrations was more convincing than providing a logical argument. She insistently claimed that arguments 4 were not convincing for her at all.

Task C was an unfamiliar case; none of the participants had experienced this type of task in their classrooms. Thus, both participants struggled with the task and neither of them could come up with a method to find the area of the figure presented. After they presented the method, their answers also differed. Even though Sara confirmed the method would work, after more thought she realized that there might be two rectangles with the same perimeter and different areas or vice versa. However, she also concluded that being able to refute argument (the method in this case) requires more than one counter example. Dacey, on the other hand, was certain that the method would work and she justified her conclusion by stating that if the outside of two shapes are equal so must the inside.

Conclusion and Discussion

The findings outline a mixed picture of what constitutes proof in the eyes of those two pre-service elementary teachers. When asked to define proof, it was clear that pre-service teachers had some experience of proof and were using this to inform their judgments about what constituted a good proof. They had experience of seeing a proof being performed and were quoting these as examples of what was required. However, despite their experience of seeing proofs in their classrooms, both participants failed to produce a proof for task A and B. This result aligns with Senk's (1989) argument that students need to be at higher levels in order to perform a proof than to be able to follow a proof. In addition, Healy and Hoyles (2000) provide evidence that students are better at choosing correct mathematical proofs than at constructing them.

In this study, even though Sara failed to apply her understanding of logical necessity to construct a proof, she was aware of the fact that inductive conclusions as in arguments 1 and 2 provide probable conclusions while, in deductive inference, the prover reaches a conclusion that is certain. Additionally, Sara exhibited different levels when she was asked to prove the statement by herself than when she was asked to evaluate different arguments constructed by others. Even though she failed to prove the statements, she recognized and selected the deductive arguments correctly.

Dacey, on the other hand, relied on examples as her primary means of justification for task A and B. She consistently justified the generalizations by stating that it worked for all the cases tested. She did not realize the limitations of such reasoning. Stylianides (2007) argues that considering empirical arguments as proof is a threat to students' opportunities to learn how to prove a proposition. Thus, one can argue that Dacey might lead students to believe that two examples would qualify as proof in her future classroom. Balacheff (1988) distinguishes between two large categories of proofs that students produced—pragmatic proofs and conceptual proofs. Pragmatic proofs are those having recourse to actual action or showings, and by contrast, conceptual proofs are those which do not involve action and rest on formulations of the properties in question and relations between them (p. 217). As in Balacheff' s definition, Dacey stated that actual action or showing was more convincing than the one which did not involve action.

This study reveals that students' levels of mathematical justification are not static. Rather, students might demonstrate different levels on different tasks depending upon their familiarity with the tasks. Fischbein (1982) argues that students choose to believe in something that seems more natural to them, subjectively, intuitively as an intrinsic property of the object. Nothing in the direct experience of the student needs such an explanation and leads to intuition. It was clear that Dacey was intuitively convinced that the outside of the shape also determines the inside in Task C. Thus she believed that if the perimeters of two shapes are equal, then the areas should be equal as well, so the method should work.

Implications and Suggestions

This study reveals that the participant PSTs do not always see the need of justifying a statement if the statement is intuitively appealing to them. In addition, they referred to the authority (the math teacher in this case) a couple of times to support their answers. Thus, it would be necessary to develop the shift of authority from teacher and textbook to the whole class. A classroom environment where mathematical ideas are not only constructed individually but also socially as students participate in meaningful activities (Cobb, Wood, Yackel, & McNeal, 1992; Yackel & Cobb, 1996) has the potential of generating mathematical justifications among prospective teachers.

I believe that a balance between visual reasoning and deductive reasoning seems to be a direction to pursue—discussing with students the role that examples play in proving statements in mathematics (Knuth, Choppin, & Bieda, 2009) while also creating learning opportunities for students to encounter both inductive and deductive proofs, so that students may develop not only a deeper understanding of proof but also a deeper understanding of the underlying reasons for using deductive proofs (Knuth, 2002).

If the goal is to help students develop a strong understanding of proof—especially in a deductive manner—teachers should assess students' current knowledge (common difficulties or misconceptions) in order to help them gradually refine their knowledge (Harel & Sowder, 2007). The framework used in this study may be a useful tool for teachers not only for assessing students' development in order to seek ways in which to help students gradually refine their perceptions but also for examining their perceptions of the nature of proof.

References

Balacheff, N. (1988). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers, and children* (pp. 216–238). London: Hodder & Stoughton.

Bell, A. W. (1976). A study of pupils' proof–explanations in mathematical situations. *Educational Studies in Mathematics, 7*(1/2), 23–40.

Burger, W. F., & Shaughnessy, M. J. (1986). Characterizing the van Hiele levels of development in geometry. *Journal for Research in Mathematics Education, 17*, 31–48.

Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics, 24*(4), 359–387.

Cobb, P., Wood, T., Yackel, E., & McNeal, B. (1992). Characteristics of classroom mathematics traditions: An interactional analysis. *American Educational Research Association, 29*(3), 573–604.

Dreyfus, T. (1999). Why Johnny can't prove. *Educational Studies in Mathematics, 38*, 85–109.

Fischbein, E. (1982). Intuition and proof. *For the Learning of Mathematics, 3*(2), 9–18.

Hanna, G. (1989) More than formal proof. *For the Learning of Mathematics, 9*(1), 20–25.

Hanna, G. (2000). Proof, explanation, and exploration: An overview. *Educational Studies in Mathematics, 44*(1/2), 5–23.

Harel, G., & Sowder, L. (2007). Toward a comprehensive perspective on proof. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 805–842). Charlotte, NC: Information Age.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31,* 396–428.

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery*. Cambridge University Press.

Knuth, E. J. (2002). Proof as a tool for learning mathematics. *Mathematics Teacher, 95*(7), 486–490.

Knuth, E. J., Choppin, J. M., & Bieda, K. N. (2009). Examples and beyond. *Mathematics Teaching in the Middle School, 15*(4), 206–211.

Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education*, 20(1), 41–51.

Morris, A. K. (2002). Mathematical reasoning: Adults' ability to make the inductive-deductive distinction. *Cognition and Instruction, 20*(1), 79–118.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Quinn, A. L. (2009). Count on number theory to inspire proof. *Mathematics Teacher, 103*(4), 298–304.

Reid, D. A. (2002). Conjectures and refutations in grade 5 mathematics. *Journal for Research in Mathematics Education, 33*, 5–29.

Senk, S. L. (1989).Van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education, 20*(3), 309–321.

Simon, M. A. (1996). Beyond inductive and deductive reasoning: The search for a sense of knowing. *Educational Studies in Mathematics, 30*(2), 197–210.

Simon, M. A., & Blume, G. W. (1996). Justification in the mathematics classroom: A study of prospective elementary teachers. *Journal of Mathematical Behavior*, *15*, 3–31.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38*, 289–321.

Weber, K., & Alcock. L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56*(2/3), 209–234.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–477.

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.